Optical Disk White Paper

RND-90-006

1990

SYSTEM DEVELOPMENT DEPARTMENT

Serge Polevitzky

1518[1]

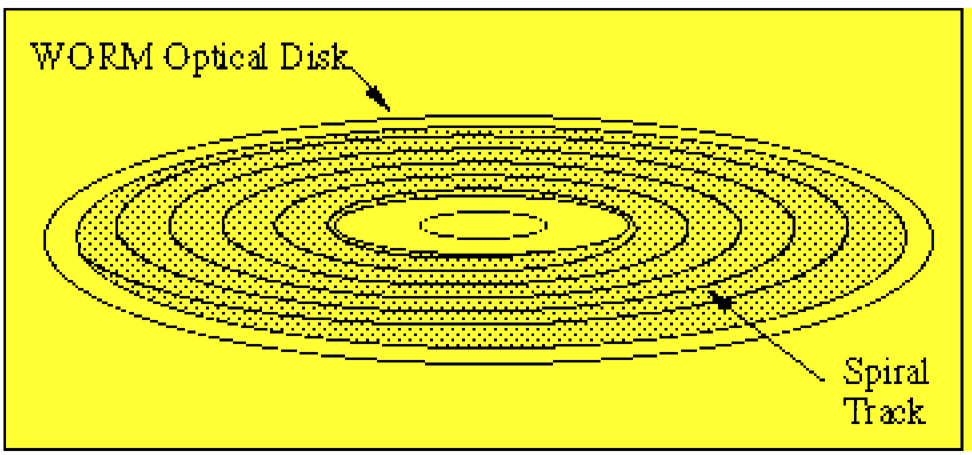


Figure 1. Continuous spiral track on WORM optical disk.

1.

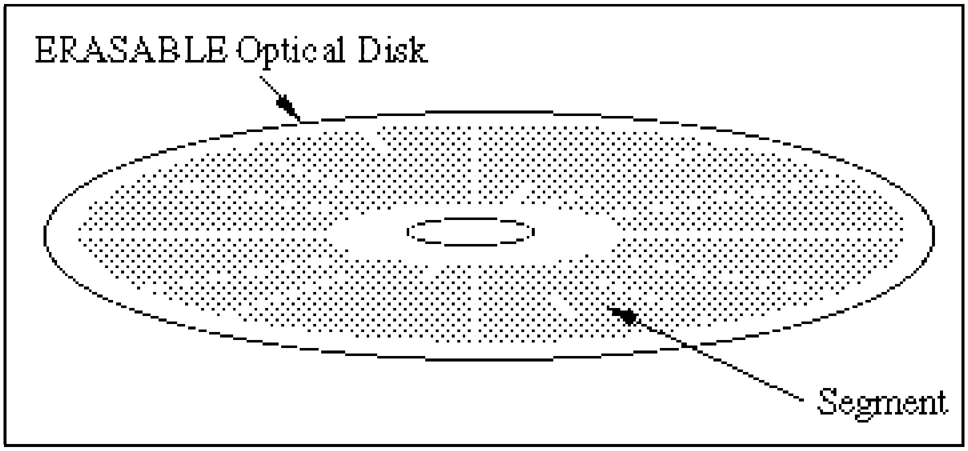


Figure 2. Segments on optical disk.

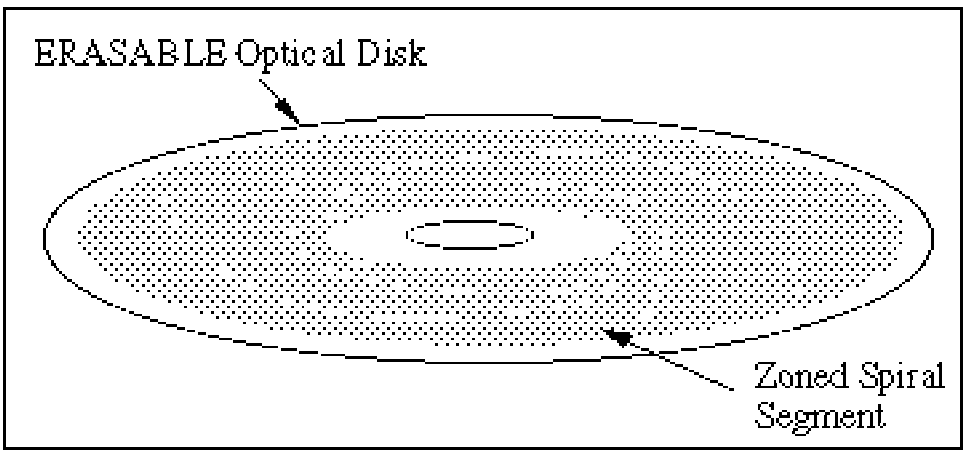


Figure 3. Maxoptix Gigabyte zoned spiral sector erasable optical disk.

WORM Performance

SIZE of MEDIA 5.25” 12”

- READS300 KBytes/sec680 KBytes/sec
- WRITES300 KBytes/sec680 KBytes/sec
- SEEKS100 millisec 300 millisec
- CAPACITY256 MBytes6.55 GBytes (single-sided)
- CAPACITY550 MBytes13.1 GBytes (dual-sided)

ERASABLE Performance

SIZE of MEDIA: 5.25” ANSI5.25” Vendor-Specific

- READS:300 KBytes/sec320 KBytes/sec
- WRITES: 75 KBytes/sec* 90 KBytes/sec*
- SEEKS:100 millisec† 100 millisec†
- CAPACITY:550 MBytes1.0 GBytes

† Exception: Maxoptix Tahiti

- SEEKS 35 millisec   35 millisec

* Estimated; vendors quote “transfer” times and give read times.

Conventional Magnetic Performance

SIZE of MEDIA	5.25” SCSI	9” IPI-2
• READS (slow)	1.2 MBytes/sec*	3 MBytes/sec†
• READS (fast)	4.0 MBytes/sec	6 MBytes/sec
• WRITES	Same	Same
• SEEKS	10 millisec	16 millisec
• CAPACITY	1.7 GBytes	2.0+ GBytes

*SCSI supports synchronous (1.25 MBytes/sec) and asynchronous

(4.0 MBytes/sec) rates;

†IPI-2 drives are delivered in 3 MBytes/sec and 6 MBytes/sec configurations;

READS and WRITES are accomplished at the same data rates.

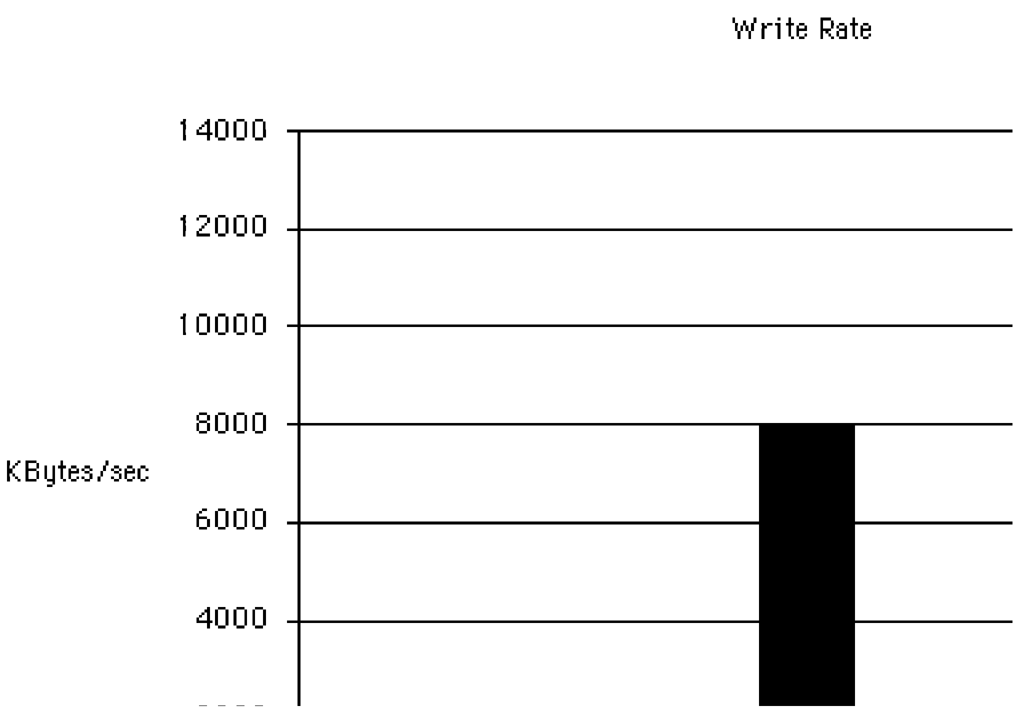


Figure 4. Comparative rates for writing data to media.

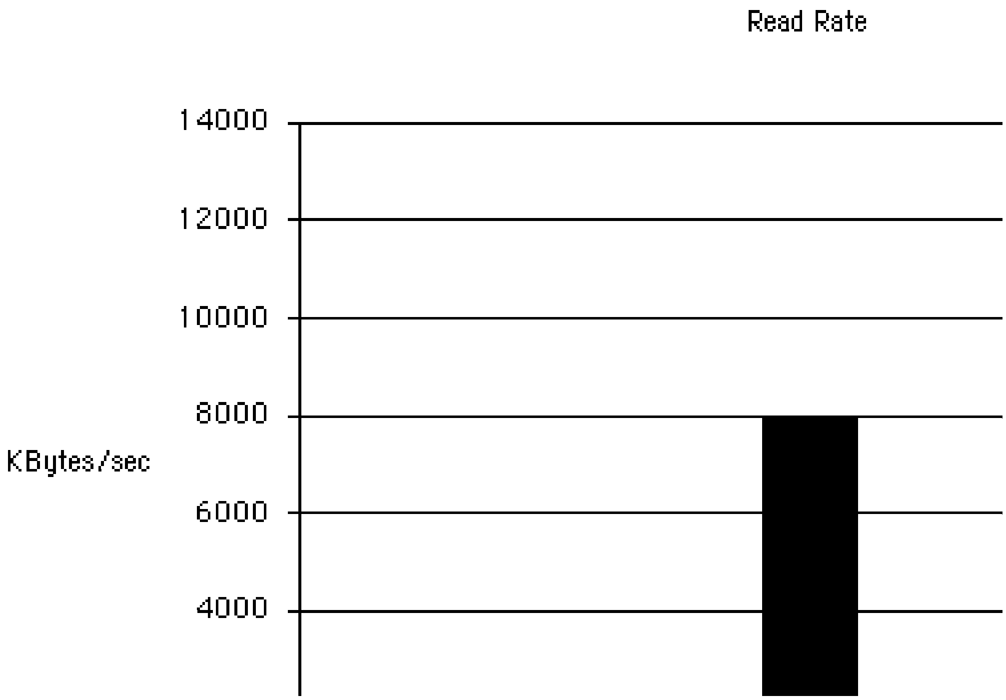


Figure 5. Comparative rates for reading data.

If a scientist uses the High Speed Processor (HSP) to create a solution file, and wants to transfer that file to a workstation for visualization and analysis, a large amount of data needs to be transferred. The following chart compares the time required for writing a 160 MByte file by WORM, ERASABLE, conventional magnetic disks (SCSI and IPI-2), 8-mm tape, ethernet, and ULTRA.

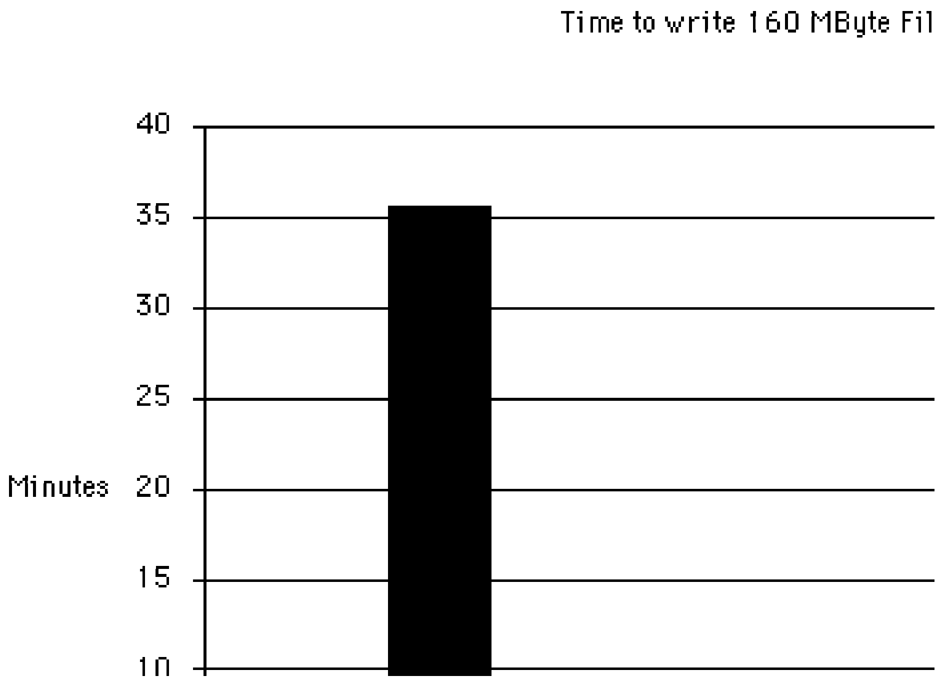
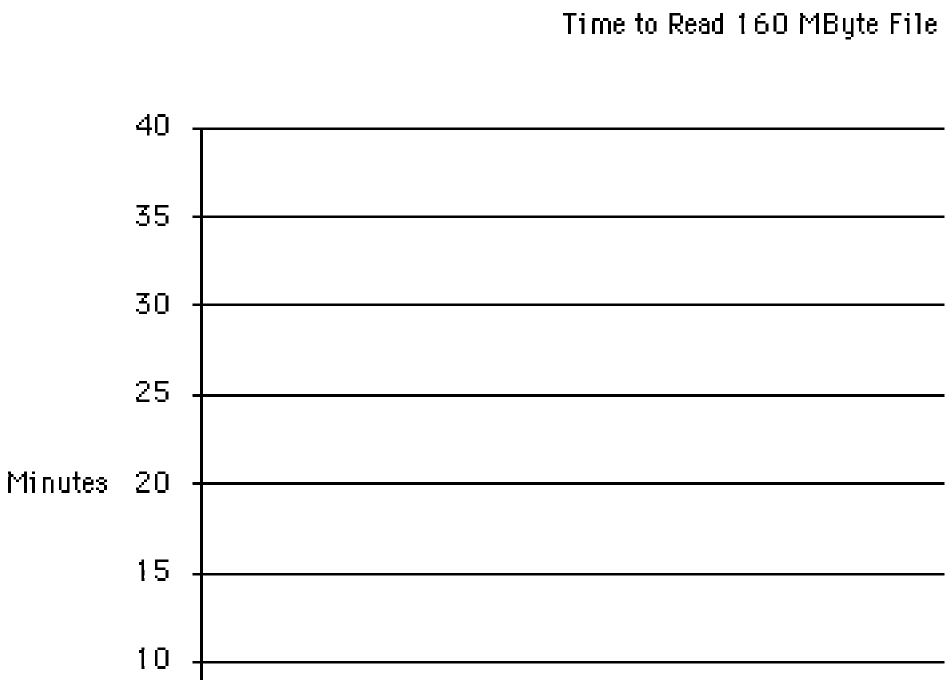


Figure 6. Comparative time needed to write a 160 MByte file.



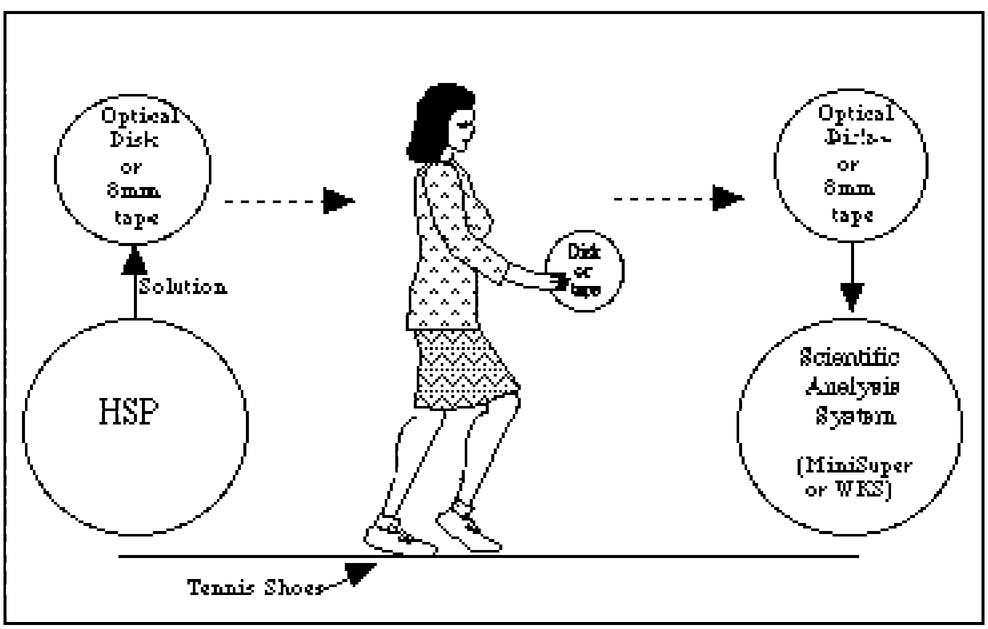


Figure 8. Manual transport of optical disk from high speed processor..

The suggestion has been made that optical disks or 8-mm tapes could be used to store the large solution files created by the HSPs. When a solution file had been written to an optical disk drive attached to the HSP, the scientist would pick up the disk or tape and carry it back to his or her workstation (or minisuper), where the solution would be visualized.

The advantages suggested by this method might be:

- The file could be removed from the Cray.
- Stress on the network would be relieved.

The disadvantages are:

- The process would be slower than a good network.
- The scientists would wind up with drawers full of disks, all vulnerable to the Coke-spill syndrome (but networks also can trash file).
- The scientists would have to take the time to master a new technology.
- Exchanging data with off-site researchers would be hampered by incompatible technologies.
- Optical disks and 8-mm tapes are not yet as cheap or fast as conventional disks.

Medium Speed LAN for users

0.533 MB/s2 $*$ 0.533 MB/s = approximately 1.06 MB/s

Time to Write a 160 MByte Solution File

MEDIA	Data Rate	Time
• WORM	680 KBytes/sec	4 minutes (235 secs)
• ERASABLE	75 KBytes/sec	35 minutes (2133 secs)
• EXAbyte 8500	500 KBytes/sec	5 minutes 20 secs (320 secs)
• ethernet	.5 MByte/sec	5 minutes 20 seconds (320 secs)
• ULTRA	2.2 MByte/sec	1 minute 12 seconds (72 secs)

WORM may be able to deliver the performance sought. Again, the delivered performance for either reading or writing may be significantly lower than the performance data quoted above.

Once again, before any commitment to a new media or technology is made, NAS should prototype several of the more-promising hardware-software offerings before creating a library of Terabytes of information.

Table 5: Optical Drive Systems available for Iris (23 October 1989)

WORM Subsystems

Vendor	Model	Interface	Capacity unformatted (Mbytes)	Avg. Seek Time (msec)	R/W Transfer Rate (Kbytes/sec)	Price Unit/Media	Price/Mbyte
ATG Gigadisc	GD1002	VME	2000	115	470/200	$12,300.00/ $550.00	$6.43
ATG Gigadisc	GD6000	VME	6400	90	990/990	$18,600.00/ $820.00	$3.03
Introl Corporation	Sterling 654W	VME	654	60	100	$8720.00/ $156.00	$13.33
Q-Systems	OFS/ SGI-WOSD	SCSI	3200		150/80	$28,500.00/ $360.00	$9.02
Aquidneck	OAS	VME	3200		100	$36,932.00/ $337.00	$11.64

Erasable Optical Subsystems

Vendor	Model	Interface	Capacity unformatted (Mbytes)	Avg. Seek Time (msec)	R/W Transfer Rate (Kbytes/sec)	Price Unit/Media	Price/Mbyte
Introl Corporation	Sterling 650E	VME	650	95	50	$10,520.00/ $236.00	$16.18
Genesis Imaging	S-501	SCSI	650	95	690	$5,995.00/ $230.00	$9.57
Q-Systems	OFS/	SCSI	650	90	300/160	$7500.00/	$11.92

SGI-MOSD
$250.00

Multidisc Subsystems

VendorModel	Interface	Capacity	Avg Access	R/W Transfer
Price	Price/Mbyte			
		unformatted	Time	Rate
Unit/Media				
		(Gigabytes)	(msec)	(Kbytes/sec)
EpochEpoch-1	NFS	31.3		300
$162,700.00	$5.20			
Systemsmodel 33		(up to 150)		
ATGGD6000	VME	704	10	990/990
$315,200.00	$0.45			
Gigadisc		(up to 912.4)		
Q-SystemsOFS/SCSI		20	3	300/160
$52,500.00	$2.62			
(erasable)SGI-MOJB				
Q-SystemsOFS/SCSI		160	3	150/80
$159,000.00	$0.99			
SGI-WOJB				
AquidneckOASVME		164		100
$140,500.00	$0.86			
Systems		(up to 1000)		
International				

Table 6: Available for Sun; drivers could be ported

(23 October 1989)

WORM Subsystems

Vendor	Model	Interface	Capacity unformatted (Mbytes)	Avg. Seek Time (msec)	R/W Transfer Rate (Kbytes/sec)	Price Unit/Media	Price/Mbyte
Delta Microsystems*	SS-622W	SCSI	622	60	269/110	$6000.00/ $150.00	$9.89
Delta Microsystems*	SS-2000WA	SCSI	2000	150	220/150	$18,750.00/ $550.00	$9.65
General Microsystems	OL/D440	VME	2400	150	334	$21,150.00/ $599.00	$9.06

Erasable Optical Subsystems

Vendor	Model	Interface	Capacity unformatted (Mbytes)	Avg. Seek Time (msec)	R/W Transfer Rate (Kbytes/sec)	Price Unit/Media	Price/Mbyte
Pinnacle Micro	REO-650	SCSI	650	65	650	$4,790.00/ $250.00	$7.75
RELAX Technology*	Erasable Optical Plus	SCSI	570 (formatted)	67	693	$3095.00/ $225.00	$5.53
Maxtor*	Tahiti I	next yr	1000	35	800/400		
SUMMUS Computer Systems	LightDisk 650	SCSI	650	90	620	$5600.00/ $250.00	$9.00
Alphatronix*	Inspire	SCSI	650	83	325	$6995.00/ $250.00	$11.15

Multidisc Subsystems

Vendor	Model	Interface	Capacity unformatted (Gigabytes)	Avg Access Time (msec)	R/W Transfer Rate (Kbytes/sec)	Price Unit/Media	Price/Mbyte
Pinnacle Micro (erasable)	REO-16000	SCSI	16	650		$63988.00	$3.99
Alphatronix*	Inspire	SCSI	30	325		$74,900.00	$2.49
Delta Microsystems*		SCSI	up to 200			$250,000.00	$1.25

* Vendors are willing to port their device driver to the Iris or are willing to give us source or detailed specifications .

A simple yardstick for estimating a disk's performance is to look at the number of sectors per track and the number of surfaces in the Head Disk Assembly, or HDA. The disk cannot read or write more sectors than pass under the read-write head(s), so if we know the number of sectors per revolution and know the spin rate, we can find an upper bound as to the number of sectors (hence, bytes) that can be read or written per revolution.

Some disks vary the number of sectors per track, packing more tracks at the edge than toward the center. So it is often possible to read and write more data per revolution at the edge than at the center. Hence, data transfer rates quoted for the "edge" can be higher than for middle or the inner tracks.

Below is an example of how the manufacturer's quoted sustained data transfer rate may be higher than the "real" sustained data rate, and the manner in which both manufacturer's and "real" were derived:

130mm (5.25") ISO/ANSI Standard MO (ERASABLE) disk —

Manufacturer's Quoted User Data Transfer Rate:
680 KBytes/sec (1024 byte/sector) — 17 sectors/tracks
620 KBytes/sec (512 byte/sector) — 31 sectors/tracks
[Source, Tecmar LaserVault™ brochure dated 5/90]

How did Tecmar come up with their stated Data Transfer Rate ? Well, a single-headed disk has 17 sectors / track at 1024 Bytes/sector. Rotational speed equals 2400 RPM (= 40 revolutions/second).

40 revolutions/sec $*$ 17 sectors/revolution $*$ 1K bytes/sector = 680KBytes/second

Or, for 512 byte sectors (31 sectors/track):

40 revolutions/sec $*$ 31 sectors/revolution $*$ 0.5K bytes/sector = 620 KBytes/second

So we can backtrack to derive the manufacturer's quoted sustained data rates. But are these what we should expect for sustained data rates for the file sizes we will be creating?

Consider that both the 620 or 680 KByte data transfer rates can be sustained for only 1 full revolution (or 1/40th of a second = 25 milliseconds). That is, either 31 or 17 sectors are transferred every rotation—every 25 milliseconds. The read-write head must then seek to the next track (hopefully, a short distance), and then wait for the rotational latency until the desired next-beginning sector appears under the heads.

A seek latency of 25 milliseconds and half the 25 milliseconds rotational latency is added to each full sector transfer. This translates into 25 + 12.5, or 37.5 milliseconds of overhead for every 25 milliseconds of "real" reading or writing. Real, sustained data throughput for a 160 MegaByte file becomes something more on the order of (1024 bytes/sector):

17 sectors transferred every 25 milliseconds (real)+ 37.5 milliseconds (overhead, to set up for the next transfer), or 17 sectors $*$ 1 KB/sector for every 25.0+37.5 milliseconds, or 17KB every 62.5ms, or multiplying by 16 to get seconds = 272KB/sec.

Remember, the vendor qouted 680 KBytes/sec for this drive.